September 1, 2015

Evaluation of Information Quality and Automation Visibility in Information Automation on the Flight Deck

Michael C. Dorneich, *Iowa State University*
Rachel Dudley, *Iowa State University*
William H. Rogers, *Honeywell International*
Emmanuel Letsu-Dake, *Honeywell International*
Stephen D. Whitlow, *Honeywell International*, et al.

Evaluation of Information Quality and Automation Visibility in Information Automation on the Flight Deck

Michael C. Dorneich[1], Rachel Dudley[2],
[1]Industrial and Manufacturing Systems Engineering, [2]Human-Computer Interaction, Iowa State University

William Rogers, Emmanuel Letsu-Dake, Stephen D. Whitlow, Michael Dillard, Erik Nelson
Human-Centered Systems, Honeywell Laboratories

An empirical study was conducted to evaluate human factors issues related to the automation visibility and information quality of an information automation system. Information automation is responsible for the collection, processing, analysis, and presentation of information to the flightcrew. Previous analysis had identified a set of characteristics particularly applicable to aircraft flight deck information automation and associated human factors issues. Air transport pilots interacted with an example information automation system in ways that allowed investigation of the information automation characteristics of automation visibility and information quality. The evaluation found that poor information quality appeared to be difficult for participants to detect, even when they were presented with the highest automation visibility level. In the times that they did not successfully compensate, participants tended to over-trust the automation, so when information was missing and they were under high workload, they chose the top plan suggested by the automation even though it was not the truly best plan. Trust in automation was reduced by low information quality, but compensated for by increased automation visibility.

INTRODUCTION

Beyond current operations, there is considerable research and development in new automation, procedures, and concepts to safely and efficiently handle an increasing demand for air travel. Next Generation Air Transportation System (NextGen) will utilize satellite-based navigation and interconnected database systems to guide and track air traffic more precisely than was previously feasible (FAA, 2013). NextGen operational concepts and technologies will dramatically affect both the types and amount of information available on flight decks (JPDO, 2009). This transformation will result in increasing automation to take advantage of the likely increase in the amount of available information (Landry, 2009). Information automation refers to automation devoted to the management and presentation of relevant information to flightcrew members (Abbott, McKenney & Railsback, 2013). Information automation processes and presents information to support a pilot's tasks and awareness and does not include cases where automation decisions and actions directly affect the aircraft performance, flightpath or systems.

The literature on human factors aspects of flight deck "automation" typically does not distinguish among different categories of automation (*e.g.* Tenney, Rogers, & Pew, 1998; Funk et al., 1999). However, there may be different human factors issues depending on the category of automation that is being considered. Work has focused on the implication of automated control of the aircraft. However, much of the automation currently being developed pertains to information support rather than control of the aircraft. Certain human factors issues and pilot errors which might be prevalent in interacting with information automation may be minimal for other types of automation. Conveying the right information at the right time to the flightcrew and accepting input from them in a user-friendly manner is critical for safe operations. The Performance-based operations Aviation Rulemaking Committee/Commercial Aviation Safety Team Flight Deck Automation Working Group has recommended that a stronger definition of information automation is needed, as well as defining terms associated with it (Abbott et al., 2013). Based on definitions of Billings (1991) and Parasuraman, Sheridan, & Wickens (2000), we developed a two-dimensional description of flight deck information automation that considered what is being managed (aircraft, systems, or information), and the stage of information processing involved (Dudley et al, 2014). A set of characteristics described a system's operation or behavior to identify potential human factors issues (Dudley et al, 2014; Rogers et al, 2015).

An empirical study was conducted to evaluate human factors issues related the interaction between Automation Visibility and Information Quality. Results led to conclusions and considerations in the assessment of information systems.

INFORMATION AUTOMATION

Information Automation Characteristics

In broad terms, information automation is the programming logic that dictates what, when, and how information is presented to the flightcrew. In previous work, nine characteristics, related to automation automation-human interaction, specific to information automation were identified (Dudley et al, 2014; see Table 1), The study focused on two characteristics: *Information Quality* and *Automation Visibility*.

Table 1. Final set of information automation charcteritcis

System Characteristics	Human-System Interaction Characteristics
Functional Complexity	Display Complexity
Information Quality	Automation Visibility
Adaptiveness	User Interaction Requirements
Level of integration	Compellingness
Degradation Behavior	

Information Quality

Previous research in *Information Quality* originated in database administration and management of information systems (Reeves & Bednar, 1994; Wang & Strong, 1996; Myers, Kappelman, & Prybutok, 1997; Pipino, Lee, & Wang, 2002; Stvilia, Gasser, Twidale, & Smith, 2007; Batini, Cappiello, Francalanci, & Maurino, 2009). In this domain, there are several attributes that play an important part in the overall concept of information quality. Wang and Strong (1996) identified four properties of high quality data: 1) intrinsically good, 2) contextually appropriate for the task, 3) clearly represented, and 4) accessible to the data consumer. They further identified 15 dimensions within the four categories to capture the *usefulness* of information as a product, or commodity, to the consumers who seek it. In our empirical evaluation, we focused on the information quality dimensions of completeness and appropriate amounts of data.

The dimensions of information quality provided by Wang and Strong (1996) are relevant to pilots as consumers of flight deck information automation output. For example, automatic de-cluttering of a display aims to provide pilots with only the most relevant and timely information for a given situation (Billings & Woods, 1994). Information quality human factors issues include: erroneous input; incomplete, uncertain, or probabilistic input; bad assumptions or factors which are not accounted for (Wand and Wang, 1996); flawed or imprecise processing (*e.g.* due to a limited or constrained model) (Ockerman and Pritchett, 2002); variety of data sources (and potential for conflicting inputs); inaccurate or unreliable outputs (Rovira et al., 2007); and untimely information.

Information Automation Visibility

Information automation visibility refers to the ability of an automation system to provide adequate feedback about its current state, what information was used, and how the information was processed (Endsley, 1996; Whitlow, Dorneich, Funk, & Miller, 2002). This characteristic may also be referred to as opacity (*e.g.* Andre & Cutler, 1998; Bisantz, Marsiglio, & Munch, 2005). In order for automation to be visible, feedback must provide a view into the automation state in a manner which can be properly interpreted by the operator (Woods, 1996) and allows the operator to predict its behavior (Scerbo, 1996). This includes information about input sources and how the system is generating outputs. High automation visibility can enhance predictability by fostering the development of an accurate mental model and enable flightcrew to verify outputs. Poor automation visibility could result in a loss of situation awareness and an increase in workload (Sarter & Woods, 1994; 1995). However, the appropriate amount and timing of automation visibility (*e.g.*, "explanatory" information) should be carefully evaluated because too much information presented at an inappropriate time (*e.g.*, during time-critical tasks) could add workload and head down time (Deveans & Kewley, 2009; Degani, Barshi, & Shafto, 2013). Automation visibility attributes may contribute to information automation human factors issues including: poor predictability of behavior (Mumaw, Sarter, & Wickens, 2001; Scerbo, 1996); hard-to-detect input or processing errors (Sarter & Woods, 1995, Sarter et al, 1997); difficult-to-assess information quality and verify outputs (Mumaw et al., 2001, Sarter et al, 1997); and inadequate, inappropriately timed, or inappropriate amount of system feedback (*e.g.*, modes) (Abbott et al., 2013; Kaber et al., 2001).

METHOD

An empirical study was conducted in a low fidelity simulator and an example information automation system.

Participants

Twelve air transport pilots from a cross section of regional and major airlines were recruited. All were male. Five were Captains and seven were First Officers. All had flown glass cockpits. Average age was 34 (range 24-56) and flight hours were about 7,000 hours (range: 2,000-14,000).

Task Application: Diversion Aid

Overview. The Diversion Aid was designed for possible NextGen decision support, when pilots might be called upon to take more responsibility for the decision of if, when, and where to divert an aircraft. The Diversion Aid integrates multiple information sources on the current state of flight, aircraft, maintenance, crew, and passenger schedules to display the implications of diversion decisions to pilots. The goal is for pilots to integrate the goals and priorities of interested airline operations stakeholders into the decision making process. This was based on a previously developed system (Dorneich et al., 2004) that uses a set of policy statements developed after conducting interviews with airline dispatchers, pilots, and various stakeholders.

Policies. A set of company policy statements was established to represent the operational priorities of all stakeholders affected by diversion decisions. These policies are used to assess the overall "goodness" of a diversion plan. Each policy was associated with cost points for each statement that is violated by a particular plan. For example, diverting a flight with an unaccompanied minor costs 10 points, while delaying a flight greater than 15 minutes costs 8 points. The policy statements are adapted from a list of policy statements developed by Dorneich et al. (2004) after conducting interviews with airline dispatchers, pilots, and various stakeholders. The goal of selecting a diversion option is to minimize the total cost incurred by the selected option: The lower the cost, the better the plan.

Displays. The Diversion Aid presented the original scheduled flight plan, followed by up to three diversion plans, based on minimizing the number of policy violations. Figure 2 depicts the graphical representations of the original plan with a ranked ordered list (best first) of diversion plan recommendations. The depiction of each individual policy violations (and their score) allows pilots to understand how the Diversion Aid ranked one plan ahead of another.

Tasks. A flight simulation was presented to the participants to help provide a sense of realism to the trials. The simulation began approximately ten minutes from top of

decent. After 60 to 90 seconds, the need for a diversion was announced and the participant was asked to make a recommendation within five minutes. The participant then started the Diversion Aid, reviewed its recommended plan(s), and decided whether to accept one of the plans or to reject its recommendation(s) if he or she could devise a better plan. The participant did not need to create a different plan. A help menu was available that displayed the set of policies

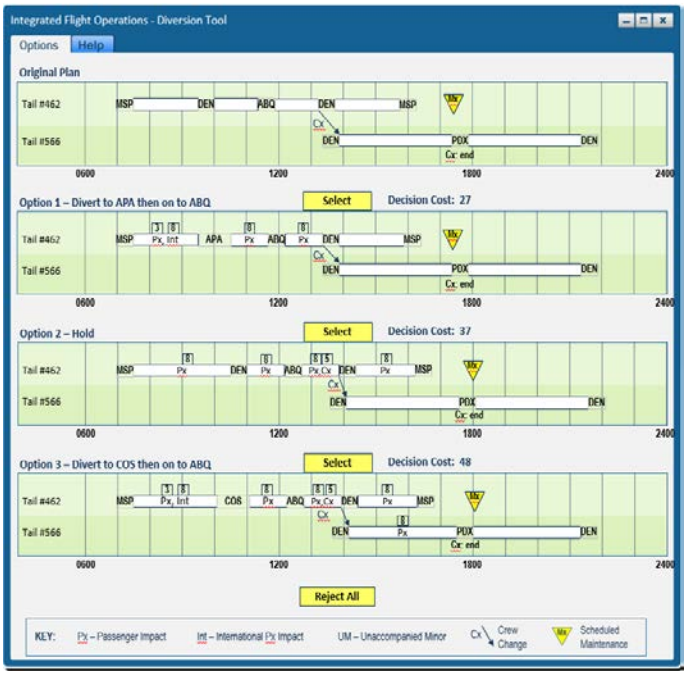

Figure 1. Diversion Aid with the top three ranked options.

Independent Variables

Table 2 describes the levels of A*utomation Visibility* (low, medium, high) and *Information Quality* (low, high).

Table 2. Independent variables for the diversion segment study.

Independent Variable	Levels	Description
Information Quality	Low	Some relevant information not included in calculation of total diversion decision cost
	High	All relevant information was included in the calculation of total diversion decision cost
Automation Visibility	Low	Best option only
	Medium	Rank-ordered list of the three best options with no policy costs shown
	High	Rank-ordered list of the three best options with policy costs shown

In the high Information Quality condition, the correct selection was always the top option on the display. In the low Information Quality condition, the correct selection was not the top selection, because the automation was missing information that resulted in incorrect scoring of the options. However, participants were given knowledge of this "missing" information during the pre-flight briefing. Thus in the low Information Quality conditions, the automation's highest ranked plan was not actually the best plan – participants were expected to recognize that a different plan was better once they included the missing information (known to them) into to their assessment. A third independent variable was originally included: *Display Modality* (text, graphic). The information content was identical between the two modes. However, since there were no statistically significant differences between the text and graphic displays, the data were collapsed over this variable.

Dependent Variables

Decision Performance was measured by the time to make a diversion decision, and correctness of the decision. Time to make a decision was the elapsed time from the start of the diversion task until participants made their diversion plan selection. The participant could accept one of the Diversion Aid plans or reject its recommendation(s). They could also reject all the plans shown if they felt that the options shown were flawed. In the medium/high Automation Visibility and low Information Quality trials, the best plan (correct selection) was listed below the automation's highest ranked plan; in the low Automation Visibility condition, only one option shown by the automation, so if the participant recognized that there was missing information, he or she could choose to reject the plan. *Workload*, *Automation Awareness* and *Trust* where measured after each trial. See Table 3.

Table 3. Dependent variables and assocaited metrics.

Dep. Variable	Metric	Frequency
Performance	Time to make a selection	once per trial
Performance	Selection of best plan	once per trial
Workload	NASA Task Load Index	once per trial
Automation Awareness	Understanding the basis of automation recommendations	once per trial
Trust	Survey question	once per trial

Experimental Design and Procedure

The experiment was a 2 (Information Quality) x 3 (Automation Visibility) within-subjects design. Scenario order was counterbalanced to avoid learning effects. Participants were trained on the Diversion Aid, tasks, and post-trial questionnaires. Participants practiced using the Diversion Aid in the simulator until they were able to make a diversion plan selection within a five minute time limit. Participants performed a total of six different diversion scenarios. After each diversion decision, they filled out the NASA-TLX workload scale and the post-trial questionnaire. After all six scenarios were completed, participants filled out a post-experiment questionnaire and were provided a short debrief.

RESULTS

A repeated measures analysis of variance was used. Results were considered significant if <= 0.05, and marginally significant if alpha > 0.05 and <= 0.1. All results charts depict means and standard error bars.

Decision Performance

Plan Selection. Participants correctly identified the best plan in 36% of the low Information Quality trials and 86% of

the high Information Quality trials. Information Quality was significant ($F_{(1,11)} = 32.98$, $p < 0.00013$). Automation Visibility was not significant.

Time to make a Selection. Automation Visibility was a significant factor ($F_{(2,22)} = 3.67$, $p < 0.042$) for this measure, with the low Automation Visibility condition being significantly faster ($t_{(22)} = 2.15$, $p < 0.043$) than the high Automation Visibility condition. Figure 1shows the time to make a selection as a function of Automation Visibility.

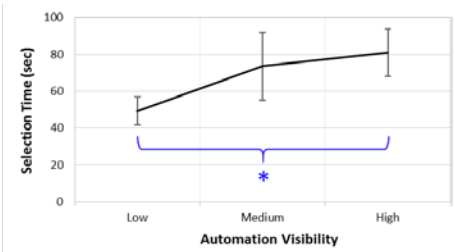

Figure 2. Selection time as a function of Automation Visibility.

Workload

Total Workload. Total workload showed no significant results for any of the dependent variables, nor their interaction.

Mental Demand: Automation Visibility was marginally significant ($F_{(2,22)} = 3.41$, $p < 0.051$). Low Automation Visibility level was significantly lower than the medium level ($t_{(22)} = 2.10$, $p < 0.047$; see Figure 2 right).

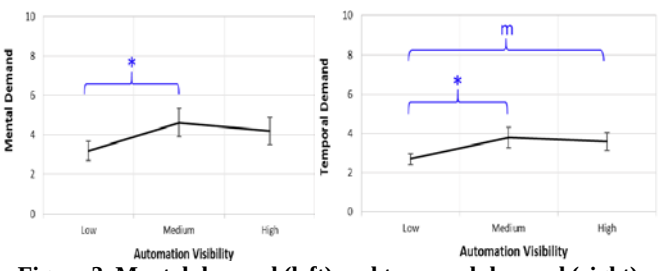

Figure 3. Mental demand (left) and temporal demand (right).

Temporal Demand: Automation Visibility was significant ($F_{(2,22)} = 4.56$, $p < 0.022$). The paired t-test showed that the low Automation Visibility level resulted in significantly lower temporal demand as compared to the medium level ($t_{(22)} = 2.38$, $p < 0.027$) and was also marginally significantly lower than the high Automation Visibility condition ($t_{(22)} = 1.92$, $p < 0.068$). Figure 2 (left) shows the results of the temporal demand as a function of Automation Visibility.

Automation Awareness

After every trial, participants were asked for their understanding of how the Diversion Aid arrived at its recommendations. Automation Visibility was significant factor for what? ($F_{(2,22)} = 5.08$, $p < 0.015$), Information Quality was marginally significant ($F_{(1,11)} = 4.11$, $p < 0.067$), and the interaction of these two independent variables was also marginally significant ($F_{(2,22)} = 2.66$, $p < 0.093$). The difference between medium and high Automation Visibility was significant ($t_{(22)} = 2.69$, $p < 0.013$) and the difference between the low and high Automation Visibility levels was marginally significant ($t_{(22)} = 1.69$, $p < 0.10$). See Figure 3

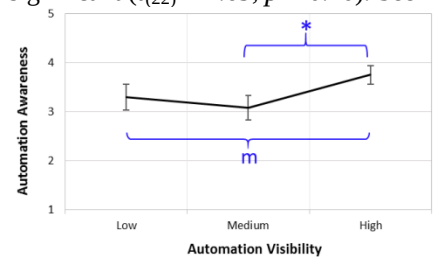

Figure 4. Automation awareness ratings.

Although the low Automation Visibility level only provided one option, participants rated their understanding of its logic closer to that of high Automation Visibility level than medium level. Having only one option presented meant that participants only had to understand one plan, rather than having to understand three plans. With the costs included in the high Automation Visibility level, the details of the logic were much more readily available.

Trust

Trust was significantly affected by the level of automation visibility ($F_{(2,20)} = 4.18$, $p < 0.030$), information quality ($F_{(1,10)} = 6.26$, $p < 0.031$;, as well as significant interaction between the two ($F_{(2,20)} = 4.15$, $p < 0.031$, respectively).

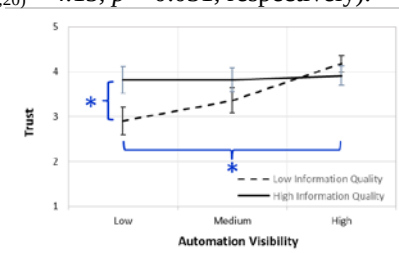

Figure 5. Trust ratings as a function of Automation Visibility.

In the high information quality, trust in the Diversion Aid did not differ due to automation visibility. In the low information quality condition, however, trust was lowest when automation visibility was low and steadily increased as automation visibility increased (and participants were more able to understand what the automation was doing).

DISCUSSION AND CONCLUSION

Participants had a difficult time detecting when the automation was reasoning over incomplete information, despite the fact that the missing information was known to them. Participants only detected the incomplete information about a third of the time. Participants tended to over-trust the automation, so when information was missing and they were under high workload, they chose the top suggested plan even though it was not the truly best plan.

In our implementation of the Diversion Aid, trust in automation was reduced by low Information Quality, but was compensated for by increased Automation Visibility. In this study, high levels of trust in automation lead to failures to override sub-optimal Diversion Aid recommendations. This had a negative impact on decision performance when

Information Quality was low. When operators were exposed to more of the Diversions Aid's automation logic, trust increased.

The level of Automation Visibility affected decision time, with low Automation Visibility leading to the fastest decision time. General consensus from the participants' qualitative responses, however, indicated that if multiple options were presented, they wanted some way to assess those options and understand the automation's reasoning (medium vs. higher Automation Visibility level). While there are advantages to making the automation's reasoning available (high information visibility), this results in a trade-off in information processing time because there is more information to review. Based on the results of the analyses and empirical studies (see also Letsu-Dake et al. (2015)), a set of recommendations and guidelines for the design and evaluation of information automation systems were developed (see Rogers et al, 2015).

ACKNOWLEDGEMENTS

This work was funded by the Federal Aviation Administration Human Factors Research and Engineering Division (contract #13-G-003, Iowa State University and Contract OTA DTFAWA-10-A-80031, Task 10, Honeywell). The opinions expressed herein are those of the authors and do not necessarily reflect the views of the FAA.

REFERENCES

Abbott, K., McKenney D., Railsback P. (2013). "Operational Use of Flight Path Management Systems", Final Report, *Performance-based operations Aviation Rulemaking Committee/Commercial Aviation Safety Team, Flight Deck Automation Working Group.*

Andre, A. D., & Cutler, H. A. (1998). Displaying uncertainty in advanced navigation systems. *In Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 42(1), 31-35. SAGE Publications.

Bass, E. J., & Pritchett, A. R. (2008). Human-automated judge learning: A methodology for examining human interaction with information analysis automation. *IEEE Systems, Man and Cybernetics, Part A:* 38(4), 759-776.

Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. ACM *Computing Survey*s (CSUR), 41(3), 16.

Billings, C. E. (1991). Toward a human-centered aircraft automation philosophy. *International Journal of Aviation Psychology*, 1(4), 261-270.

Billings, C. E. (1997). *Aviation automation*: The search for a human-centered approach. Lawrence Erlbaum Associates, Mahwah, NJ.

Billings, C. E., & Woods, D. D. (1994). Concerns about adaptive automation in aviation systems. *Human performance in automated systems: Current research and trends*, 264-269.

Bisantz, A.M, Marsiglio, S.S., & Munch, J. (2005). Displaying uncertainty: Investigating the effects of display format and specificity. *The Journal of the Human Factors and Ergonomics Society*, 47(4), 777-796.

Degani, A., Barshi, I., & Shafto, M. G. (2013). Information Organization in the Airline Cockpit Lessons From Flight 236. *Journal of Cognitive Engineering and Decision Making*, 7(4), 330-352.

Deveans, T., & Kewley, R. (2009). Overcoming information overload in the cockpit (*ORCEN Tech. Rep. No. DSE-TR-0904*). West Point, NY: U.S. Army.

Dorneich, M.C., Whitlow, S.D., Miller, C.A., & Allen, J.A (2004). "A Superior Tool for Airline Operations", *Ergonomics in Design*. Vol. 12, No. 2, pp. 18-23.

Dudley, R., Dorneich, M.C., Letsu-Dake, E., Rogers, W., Whitlow, S.D., Dillard, M., & Nelson, E. (2014). "Characterization of Information Automation on the Flight Deck," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* . Chicago, IL, October 27–31.

Endsley, M. R. (1996). Automation and situation awareness. *Automation and human performance: Theory and applications*, 163-181.

FAA. (2013). NextGen Implementation Plan. Available: http://www.faa.gov/nextgen/implementation/media/NextGen_Implementation_Plan_2013.pdf *(accessed August 13, 2013)*.

Funk, K., Lyall, B., Wilson, J., Vint, R., Niemczyk, M., Suroteguh, C., & and Owen, G. (1999). Flight Deck Automation Issues, *International Journal of Aviation Psychology*, Vol. 9, No. 2, pp. 109-123.

JPDO (2009). Concept of Operations for the Next Generation Air Transportation System. Version 3.0, Washington, DC, 153.

Kaber, D. B., Riley, J., Tan, K., & Endsley, M. R. (2001). On the design of adaptive automation for complex systems. *International Journal of Cognitive Ergonomics*, 5, 37–57.

Landry, S.J. (2009). "Flight Deck Automation", *Springer Handbook of Aviation* (Nof, ed.), Dordrecht: Springer. Chapter 68, pp. 1213-1240.

Letsu-Dake, E., Rogers, W., Whitlow, S.D., Nelson, E., Dillard, M., Dorneich, M.C., & Dudley, R. (2015). "light Deck Information Automation: A Human-In-The Loop In-Trail Procedure Simulation Study," 34th Digital Avionics Systems Conference. Prague, Czech Republic, September 13-17.

Myers, B. L., Kappelman, L. A., & Prybutok, V. R. (1997). A comprehensive model for assessing the quality and productivity of the information systems function: toward a theory for information systems assessment. *Information Resources Management Journal* (IRMJ), 10(1), 6-26.

Mumaw, R. J., Sarter, N. B., & Wickens, C. D. (2001). Analysis of pilots' monitoring and performance on an automated flight deck. In Proceedings of the Eleventh International Symposium in Aviation Psychology (p. 6). Columbus: Ohio State University.

Ockerman, J. J., & Pritchett, A. R. (2002, September). Impact of Contextual Information on Automation Brittleness. I*n Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 46, No. 3, pp. 382-386). SAGE Publications.

Parasuraman, R., Sheridan, T.B., & Wickens, C.D. (2000). A model for types and levels of human interaction with automation. *IEEE Systems, Man and Cybernetics, Part A*, 30(3), 286-297.

Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, 45(4), 211-218.

Reeves, C. A., & Bednar, D. A. (1994). Defining quality: alternatives and implications. *Academy of management Review*, 19(3), 419-445.

Rogers, W., Letsu-Dake, E., Dorneich, M., Dudley, R., Nelson, E., Dillard, M., Whitlow, S., Almquist, J., Hamblin, C, * Ott, J. (2015). "Flight Deck Information Automation: Analysis and Recommendations", Final Report to Federal Aviation Administration.

Rovira, E., Zinni, M., & Parasuraman, R. (2002, September). Effects of information and decision automation on multi-task performance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 46, No. 3, pp. 327-331).

Sarter, N. B., & Woods, D. D. (1994, March). Decomposing automation: Autonomy, authority, observability, and perceived animacy. Paper presented at the *First Automation Technology and Human Performance Conference*, Washington, DC.

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors*, 37, 5–19.

Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (2nd ed., pp. 1926-1943). New York: Wiley.

Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 37–63). Mahwah, NJ: Lawrence Erlbaum.

Stvilia, B., Gasser, L., Twidale, M. B., & Smith, L. C. (2007). A framework for information quality assessment. *Journal of the American Society for Information Science and Technology*, 58(12), 1720-1733.

Tenney, Y. J., Rogers, W. H., & Pew, R.W. (1998). Pilot Opinions on Cockpit Automation Issues. *International Journal of Aviation Psychology*, Vol. 8, No. 2, pp. 103-120.

Wand, Y. and Wang, R. (1996). Anchoring data quality dimensions in ontological foundations. Comm. *ACM* 39, 11.

Wang, R. & Strong, D. (1996). "Beyond Accuracy: What Data Quality Means to Data Consumers." *Journal of Management Information Systems*, 12(4), p. 5-34.

Whitlow, S. D., Dorneich, M. C., Funk, H. B., & Miller, C. A. (2002). Providing appropriate situation awareness within a mixed-initiative control system. *IEEE Systems, Man and Cybernetics Conference* (Vol. 5, pp. 5).

Woods, D. D. (1996). Decomposing automation: Apparent simplicity, real complexity. *Automation and human performance: Theory and applications*, 3-17.